# 사용자의 건강 상태 예측을 위한 건강 검진 데이터 패턴 분석

2020.10.22(목)~2020.11.10(화) / 박소현, 신미용, 정재민

**개요**

* 본 프로젝트는 건강검진 데이터의 각 요소들의 패턴화를 통해 추이를 분석하여 사용자가 본인의 데이터를 기입하였을 시 대응하는 추이에 따라 향후 건강 상태를 예측하기 위한 데이터 분석이다.

**목표**

1. 데이터 분류, 패턴화를 통한 각 요소 별 추이 분석
2. 사용자의 향후 건강 상태 예측을 위한 Dataset 구축

**데이터**

* 투비콘에서 건강 검진 데이터 수령

(건강 검진을 10년이상 받아온 사람들의 데이터로 column은 1069, row는 29의 xlsx 데이터)

* 건강 검진 데이터 중 일반 검진에 해당하는 항목만 사용

(고유값, 성별, 생년, 검진년도, 키, 체중, BMI, 시력(좌/우), 청력(좌/우), 혈압(수축기/이완기), 요단백, 헤모글로빈, 공복혈당, 혈청크레아티닌, 신사구체여과율, AST, ALT, GTP, 폐결핵흉부질환)

**일정**

**10.22(목)**

* 주제 선정

(사용자의 건강 상태 예측을 위한 건강 검진 데이터 패턴 분석)

**10.26(월)**

* 기획서 작성

**10.27(화) – 10.30(금)**

* + - 데이터 수령
    - 데이터 전처리

(건강 검진 데이터 중 일반 검진에 해당하는 항목만 추출)

(건강 검진 데이터의 생년, 검진년도를 이용하여 검진 시 연령 row를 제작)

(성별을 기준으로 분류)

* + - 데이터 시각화

(검진 시 연령 row를 x축으로 삼고 각 요소 시각화)

**10.31(토) – 11.06(금)**

* + - 데이터 패턴 분석

**11.07(토) – 11.10(화)**

* + - 발표 자료 제작